DETAILED ACTION

The final rejection mailed on 4/29/08 has been replaced with the current supplemental non-final Office Action to reject claims 27-29 under 35 U.S.C.101.

Response to Arguments

1. Applicant's arguments filed 2/1/08 have been fully considered but they are not persuasive.

1. Regarding lines 6-9 on page 2 of applicant's remarks, applicant states that the combination of Lee and Linzer are not combinable and the combination has "fundamental problems". The examiner respectfully disagrees. The examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See In re Fine, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and In re Jones, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality, as suggested in Linzer's column 4, lines 39-42.

The test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have

suggested to those of ordinary skill in the art. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981).

One cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); In re Merck & Co., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

It is noted that the features upon which applicant relies (i.e., portion) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See In re Van Geuns, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Regarding lines 4-5 on page 3 of applicant's remarks, applicant asserts that the term “portion” is defined differently from the Lee reference. The examiner respectfully disagrees. The term “portion” can be reasonably interpreted as a “part” or “limited quantity of anything”, as defined in Webster's Dictionary, as well as the general definition of the term “portion”. In column 42, lines 47-61, Lee discloses that each video object has an arbitrary shape, and that each video object is predefined according to its shape. Thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask. In figure 27A, Lee discloses there are at least two video portions or “parts”, wherein elements 972, 974, 976, 978, 980 and 982 are the “portions” or “parts”, in that there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970. In figure

35, Lee discloses that frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b.

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, in column 6, lines 56-58, Linzer's figure 3 discloses that elements 32-1 to 32-n are the plural MPEG encoders, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n. In other words, Linzer discloses the implementation of compressing multiple video sources with multiple temporal frame portions in that each source consists multiple GOPs with multiple frames. Thus, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content. Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality, as suggested in Linzer's column 4, lines 39-42.

Regarding lines 3-4 on page 4 of applicant's remarks, applicant asserts that the reason to combine Lee and Linzer is incomplete and insufficient. The examiner respectfully disagrees. The examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so

found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See In re Fine, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and In re Jones, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality, as suggested in Linzer's column 4, lines 39-42.

The test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981).

One cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); In re Merck & Co., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

It is noted that the features upon which applicant relies (i.e., portion) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See In re Van Geuns, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Regarding lines 4-9 on page 4 of applicant's remarks about the term "portion ", applicant argues the definition of "portion" and disagrees with the use of the term "portion". The examiner respectfully disagrees. The term "portion" can be reasonably interpreted as a "part" or "limited quantity of anything", as defined in Webster's Dictionary, as well as the general definition of the term "portion". In column 6, lines 56-58, Linzer's figure 3 discloses that elements 32-1 to 32-n are the plural MPEG encoders, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n. Linzer discloses the implementation of compressing multiple video sources with multiple temporal frame portions in that each source consists multiple GOPs with multiple frames. Thus, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content.

Regarding lines 10-12 and lines 19-22 on page 4, and lines12-15 on page 5 of applicant's remarks, applicant states that the modification of the prior art invention would be unsatisfactory for its intended purpose, and that there is no suggestion or motivation to combine the Lee and Linzer, and that the combination is "unworkable". The examiner respectfully disagrees. The examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so

found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See In re Fine, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and In re Jones, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality, as suggested in Linzer's column 4, lines 39-42.

The test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981).

One cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); In re Merck & Co., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

It is noted that the features upon which applicant relies (i.e., portion) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See In re Van Geuns, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Regarding lines 5-7 on page 5 of applicant's remarks, applicant states that the term "portion" is changed. The term "portion" is and can be reasonably interpreted as a "part" or "limited quantity of anything", as defined in Webster's Dictionary, as well as the general definition of the term "portion". As stated above and in the rejection below, Lee discloses "video portions". Lee does not disclose the "portions" comprise a "temporal, multiframe segment of the video content". However, in column 6, lines 56-58, Linzer's figure 3 discloses that elements 32-1 to 32-n are the plural MPEG encoders, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n. In other words, Linzer discloses the implementation of compressing multiple video sources with multiple temporal frame portions in that each source consists multiple GOPs with multiple frames. Thus, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content. Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality, as suggested in Linzer's column 4, lines 39-42.

2. Regarding lines 16-21 on page 5 of applicant's remarks, applicant states that the combination of Lee and Linzer would require "reconstruction" and "redesign". The examiner respectfully disagrees. In response to applicant's argument that the

examiner's conclusion of obviousness is based upon reasoning of “reconstruction and redesign”, it must be recognized that any judgment on obviousness is in a sense necessarily a reconstruction based upon hindsight reasoning. But so long as it takes into account only knowledge which was within the level of ordinary skill at the time the claimed invention was made, and does not include knowledge gleaned only from the applicant's disclosure, such a reconstruction is proper. See In re McLaughlin, 443 F.2d 1392, 170 USPQ 209 (CCPA 1971). The test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981).

Dependent claims 2-5, 7 and 8 are rejected for at least similar reasons as claim 1. Independent claim 9 and claims 10-12 are rejected for at least similar reasons as explained above for claim 1. Claim 13-14 is rejected for at least similar reasons as claim 1. Claims 15-16 are rejected for at least similar reasons as explained above for claim 1. Claims 18-20 are rejected for at least similar reasons as stated above for claim 1. Claims 21-22 are rejected for at least similar reasons as claim 1. Claims 27-35 are rejected for at least similar reasons as set forth for claims 1, 9, 13, 15, 18 and 21.

Thus, the rejection is maintained.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

The USPTO "Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility" (Official Gazette notice of 22 November 2005), Annex IV, reads as follows:

> Claims that recite nothing but the physical characteristics of a form of energy, such as a frequency, voltage, or the strength of a magnetic field, define energy or magnetism, per se, and as such are nonstatutory natural phenomena. O'Reilly, 56 U.S. (15 How.) at 112-14. Moreover, it does not appear that a claim reciting a signal encoded with functional descriptive material falls within any of the categories of patentable subject matter set forth in Sec. 101.
>
> … a signal does not fall within one of the four statutory classes of Sec. 101.
>
> … signal claims are ineligible for patent protection because they do not fall within any of the four statutory classes of Sec. 101.

Claim 27-29 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter as follows. Claims 27-29 define "**a coded bitstream having portions of the bitstream encoded using different encoders according to encoder models…**" with descriptive material. While "functional descriptive material" may be claimed as a statutory product (i.e., a "manufacture") when embodied on a tangible computer readable medium, "**a coded bitstream**…" embodying that same functional descriptive material is neither a process nor a product (i.e., a tangible "thing") and therefore does not fall within one of the four statutory classes of § 101. Rather, "signal" is a form of energy, in the absence of any physical structure or tangible material.

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1-5, 7-16, 18-22 and 27-35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lee (5,748,789) in view of Linzer (6,094,457).

Regarding claim 1, Lee discloses a method of encoding video content, the method comprising:

assigning a predefined model to each of at least two video content portions of the video content (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b); and

routing each of the at least two video content portions to one of a plurality of encoders based on a respective one of the predefined models assigned to each of the at least two video content portions (col.42, ln.47-61, Lee discloses that each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is routed or assigned a predefined encoder model by a mask of alpha values or a binary mask),

wherein the assigning a predefined model to each of the at least two video content portions (col.42, ln.47-61; note each video object has an arbitrary shape, and

that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b) further comprises:

comparing descriptors associated with each of the at least two video content portions with corresponding stored model descriptors from a plurality of predefined content models (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions), and

assigning each of the at least two video content portions to a respective best content model from the plurality of predefined content models based on the comparing of the descriptors (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe

segment of the video content (fig.3, Linzer discloses that elements 32-1 to 32-n are the plural MPEG encoders, as disclosed in col.6, ln.56-58, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n). Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality (Linzer col.4, ln.39-42).

Regarding claim 2, Lee discloses the at least two video content portions are video segments (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b).

Regarding claim 3, Lee discloses the at least two video content portions are video subsegments (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a

binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b).

Regarding claim 4, Lee discloses the at least two video content portions are video regions of interest (fig.33, element 1502, col.42, ln.34-46, and fig.35, note video object information is extracted and segmented from the input video sequence, and segments and subsegments of the regions of interest are identified, and in fig.35 discloses extracting multiple video objects 1540, 1542 and 1544b; fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b).

Regarding claim 5, Lee discloses a generic encoder model (fig.33 and col.42, ln.62-65; note object coders 1504-1508 encode video portions associated with the generic model; and fig.36, note the coder shown is used to encode the video portions).

Regarding claim 7, Lee discloses one of the plurality of predefined content models includes a generic video content model (fig.33 and col.42, ln.62-65; note object coders 1504-1508 encode video portions associated with the generic model; and fig.36, note the coder shown is used to encode the video portions).

Regarding claim 8, Lee discloses wherein assigning a predefined model to each of at least two video content portions of the video content further comprises assigning the generic video content model to a video content portion of the at least two video content portions if none of the other models from the plurality of predefined content

models is assigned to the video content portion (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b;).

Regarding claim 9, Lee discloses a method of encoding video content, the method comprising:

identifying video subsegments and regions of interest within at least two video portions from the video content (fig.33, element 1502, col.42, ln.34-46, and fig.35, note video object information is extracted and segmented from the input video sequence, and segments and subsegments of the regions of interest are identified, and in fig.35 discloses extracting multiple video objects 1540, 1542 and 1544b; fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b);

assigning a predefined encoder model to each at least two video portion according to a characteristic of each of the at least two video portions, the predefined encoder model being chosen from a plurality of predefined models or a generic model (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is

assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b);

encoding each of the at least two video portions associated with the generic encoder model with a generic encoder (fig.33 and col.42, ln.62-65; note object coders 1504-1508 encode video portions associated with the generic model; and fig.36, note the coder shown is used to encode the video portions); and

encoding each of the at least two video portions associated with the plurality of predefined encoder models with an encoder chosen from a plurality of encoders, each of the plurality of encoders being associated with one of the plurality of predefined models (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions), wherein

the assigning a predefined encoder model to each of the at least two video portions according to a characteristic of each of the at least two video portions (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements

972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b) further comprises:

comparing first descriptors associated with the at least two video portions and second descriptors associated with the subsegments and the regions of interest with corresponding stored model descriptors from a plurality of predefined content models (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions), and

assigning each of the at least two video content portions to a respective best content model based on the comparing of the first and the second descriptors (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content (fig.3, Linzer discloses that elements 32-1 to 32-n are the plural MPEG encoders, as disclosed in col.6, ln.56-58, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n).

Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding

multiple video streaming image data while maintaining high image quality (Linzer col.4, ln.39-42).

Regarding claim 10, Lee discloses producing the first descriptors associated with the at least two video portions of the video content (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions); producing the second descriptors associated with the video subsegments and the regions of interest (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions).

Regarding claim 11, Lee discloses encoding the first and second descriptors (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions, that includes coding the first and second descriptors).

Regarding claim 12, Lee discloses wherein the first and second descriptors are used to determine whether the generic encoder or an encoder from a plurality of encoders was used to encode the at least two video portions (fig.33 and col.43, ln.10-

15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions, that includes coding the first and second descriptors).

Regarding claim 13, Lee discloses a method of encoding video content, the method comprising:

if a video portion of at least two video portions of the video content relates to one of a plurality of predefined encoder models (fig.33, element 1502, col.42, ln.34-46, and fig.35, note video object information is extracted and segmented from the input video sequence, and segments and subsegments of the regions of interest are identified, and in fig.35 discloses extracting multiple video objects 1540, 1542 and 1544b; fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b),

assigning the video content portion to a related, predefined encoder model chosen from the plurality of predefined encoder models (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b);

if a video content portion of the at least two video content portions of the video content does not relate to one of the plurality of predefined encoder models, assigning the video content portion to a generic encoder model (fig.33 and col.42, ln.62-65, Lee discloses the object coders 1504-1508 are used to encode the video portions associated with the generic model, in fig.36, the coder shown is used to encode the video portions in a generic manner or model);

encoding each of the at least two video content portions associated with the generic encoder model using a generic encoder (fig.33 and col.42, ln.62-65; note object coders 1504-1508 encode video portions associated with the generic model; and fig.36, note the coder shown is used to encode the video portions in a generic manner or model); and

encoding each of the at least two video content portions associated with one of the predefined encoder models with an encoder from a plurality of encoders (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions),

wherein the assigning the video content portion to a related, predefined encoder model chosen from the plurality of predefined encoder models (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are

triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b) further comprises:

comparing descriptors associated with the video content portion with corresponding stored model descriptors from a plurality of predefined encoder models (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions), and

assigning the video content portion to a best encoder model from the plurality of predefined encoder models based on the comparing of the descriptors (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content (fig.3, Linzer discloses that elements 32-1 to 32-n are the plural MPEG encoders, as disclosed in col.6, ln.56-58, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal,

multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n). Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality (Linzer col.4, ln.39-42).

Regarding claim 14, Lee discloses wherein each encoder from a plurality of encoders is associated with one of the predefined encoder models of the plurality of predefined encoder models (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions).

Regarding claim 15, Lee discloses a method of encoding video content divided into a at least two portions (fig.33, element 1502, col.42, ln.34-46, and fig.35, note video object information is extracted and segmented from the input video sequence, and segments and subsegments of the regions of interest are identified, and in fig.35 discloses extracting multiple video objects 1540, 1542 and 1544b; fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b), each of the at least two portions being associated with either a generic encoder model or an encoder model chosen from a plurality of predefined encoder models (fig.33 and col.42, ln.62-65; note object coders 1504-1508 encode video portions associated with the generic model;

and fig.36, note the coder shown is used to encode the video portions), the method comprising:

comparing descriptors associated with the at least two portions with corresponding stored model descriptors from a plurality of predefined encoder models (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions);

assigning each of the at least two portions to a respective best encoder model from the plurality of predefined encoder models based on the comparing of the descriptors (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data);

routing each of the at least two portions that is not assigned a respective best encoder model from the plurality of encoder models to a generic encoder (fig.33 and col.42, ln.62-65, Lee discloses the object coders 1504-1508 are used to encode the video portions associated with the generic model, in fig.36, the coder shown is used to encode the video portions in a generic manner or model); and

routing each of the at least two portions assigned to the respective best encoder model of the plurality of predefined encoder models to an encoder associated with the respective best encoder model (col.50, ln.27-37, the error computed from the inter-

frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content (fig.3, Linzer discloses that elements 32-1 to 32-n are the plural MPEG encoders, as disclosed in col.6, ln.56-58, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n). Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality (Linzer col.4, ln.39-42).

Regarding claim 16, Lee discloses wherein each encoder from a plurality of encoders is optimized for each predefined encoder model of the plurality of encoder models (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions, thus optimizing the encoders for each predefined model of plural encoder models).

Regarding claim 18, Lee discloses a method of producing a bitstream coded according to video content, the method comprising:

associating each of at least two portions of the video content to either a generic encoder model or a predefined encoder model chosen from a plurality of predefined encoder models (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b);

routing each of the at least two portions associated with the generic encoder model to a generic encoder (fig.33 and col.42, ln.62-65; note object coders 1504-1508 encode video portions associated with the generic model; and fig.36, note the coder shown is used to encode the video portions in a generic manner or model); and

routing each of the at least two portions associated with an encoder model of the plurality of predefined encoder models to one of a plurality of encoders, wherein each encoder of the plurality of encoders is associated with one of the predefined encoder models (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions),

wherein the associating each of the at least two portions of the video content to either a generic encoder model or a predefined encoder model chosen from a plurality of predefined encoder models (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b) further comprises:

comparing descriptors associated with each of the at least two portions with corresponding stored model descriptors from the plurality of predefined encoder models (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions), and

associating each of the at least two portions with a respective best encoder model from the plurality of predefined encoder models or the generic encoder model based on the comparing of the descriptors (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content (fig.3, Linzer discloses that elements 32-1 to 32-n are the plural MPEG encoders, as disclosed in col.6, ln.56-58, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n). Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality (Linzer col.4, ln.39-42).

Regarding claim 19, Lee discloses multiplexing each portion and transmitting each portion in a bitstream (fig.33 and col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural video object encoders 1504-1508; and fig.36, note the coder shown is used to encode the video portions).

Regarding claim 20, Lee discloses locating subsegments and regions of interest in the extracted portions (fig.33, element 1502, col.42, ln.34-46, and fig.35, note video object information is extracted and segmented from the input video sequence, and segments and subsegments of the regions of interest are identified, and in fig.35 discloses extracting multiple video objects 1540, 1542 and 1544b; fig.27A, note there

are at least two video portions, elements 972, 974, 976, 978, 980 and 982; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b).

Regarding claim 21, Lee discloses a method of encoding a bitstream using a plurality of encoders, the method comprising:

mapping each of at least two segments extracted from video content to a predefined encoder model (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b); and

routing the at least two extracted and mapped segments to one of the plurality of encoders based on the mapping to the respective predefined encoder model (col.42, ln.47-61, Lee discloses that each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is routed or assigned a predefined encoder model by a mask of alpha values or a binary mask),

wherein the mapping each of at least two segments extracted from the video content to a predefined encoder model (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus,

each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b) further comprises:

comparing descriptors associated with each of the at least two extracted segments with corresponding stored model descriptors from the plurality of predefined encoder models (col.51, ln.4-59; note there are plural flags that can aid the determination of the video portions of the video content; col.50, ln.18-41, Lee discloses the comparison of the frames, in particular, the comparison is done with the shape of the first frame that contains its respective video portions and the shape of the second frame that contains its respective video portions), and

mapping each of the at least two extracted segments to a respective best encoder model from the plurality of predefined encoder models based on the comparing (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Lee does not specifically disclose wherein each of the at least two video portions comprise a temporal, multiframe segment of the video content. However, Linzer teaches that each of the at least two video portions comprise a temporal, multiframe segment of the video content (fig.3, Linzer discloses that elements 32-1 to 32-n are the

plural MPEG encoders, as disclosed in col.6, ln.56-58, in that each of the encoders 32-1 to 32-n compress a temporal, multiframe segment of the available video content known in MPEG as a group of frames (GOPs) or a group of frames organized in a temporal, multiframe or grouped frames that can be partitioned into multiple frame segments, and that the multiple frame segments are compressed by encoders 32-1 to 32-n). Therefore, it would have been obvious to one of ordinary skill in the art to combine the teachings of Linzer into the system of Lee for permitting accurately, efficiently encoding multiple video streaming image data while maintaining high image quality (Linzer col.4, ln.39-42).

Regarding claim 22, Lee discloses locating subsegments and regions of interest in the extracted segments (fig.33, element 1502, col.42, ln.34-46, and fig.35, note video object information is extracted and segmented from the input video sequence, and segments and subsegments of the regions of interest are identified, and in fig.35 discloses extracting multiple video objects 1540, 1542 and 1544b; fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b).

Regarding claims 27-29, Lee discloses a coded bitstream having portions of the bitstream encoded using different encoders according to encoder models associated with a subject matter of each portion of the bitstream, the coded bitstream encoded according to the method of claims 1, 18 and 21, respectively (fig.33 and col.42, ln.62-65; note different video object coders 1504-1508 encode video portions associated with

the generic model; col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural different video object encoders 1504-1508).

Regarding claim 30, Lee discloses wherein the assigning a predefined model to each of at least two video content portions of the video content portions of the video content further comprises assigning a different predefined model to each of the at least two video content portions of the video content (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b; fig.33 and col.42, ln.62-65; note different video object coders 1504-1508 encode video portions associated with the generic model; col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural different video object encoders 1504-1508).

Regarding claim 31, Lee discloses wherein the assigning a predefined encoder model to each of at least two video portions according to a characteristic of each of the at least two video further comprises assigning a different predefined encoder model to each of the at least two video portions of the video content (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder

model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b; fig.33 and col.42, ln.62-65; note different video object coders 1504-1508 encode video portions associated with the generic model; col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural different video object encoders 1504-1508).

Regarding claim 32, Lee discloses wherein the assigning the video content portion to a related, predefined encoder model chosen from the plurality of predefined encoder models further comprises assigning each of the at least two video content portions of the video content to a different one of the predefined encoder models (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b; fig.33 and col.42, ln.62-65; note different video object coders 1504-1508 encode video portions associated with the generic model; col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural different video object encoders 1504-1508).

Regarding claim 33, Lee discloses wherein assigning each of the at least two portions to a respective best encoder model from the plurality of predefined encoder models based on the comparing of the descriptors further comprises assigning each of the at least two portions to a different one of the plurality of predefined encoder models (col.50, ln.27-37, the error computed from the inter-frame shape coding is then applied to assign the best content model based on the interframe comparison of the shapes between the first and second frame data).

Regarding claim 34, Lee discloses the associating each of the at least two portions of the video content to either a generic encoder model or a predefined encoder model further comprises associating each of the at least two portions of the video content to a different encoder model chosen from the generic encoder model of the plurality of predefined encoder models (col.42, ln.47-61; note each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is assigned a predefined encoder model by a mask of alpha values or a binary mask; in fig.27A, note there are at least two video portions, elements 972, 974, 976, 978, 980 and 982, where there are triangular portions that consist of each of elements 972, 974, 976, 978, 980 and 982 to form a model of a person 970; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b; fig.33 and col.42, ln.62-65; note different video object coders 1504-1508 encode video portions associated with the generic model; col.43, ln.10-15; note the multiplexer

1510 is used to multiplex and encode video portions from plural different video object encoders 1504-1508).

Regarding claim 35, Lee discloses wherein the mapping each of at least two segments extracted from the video content to a predefined encoder model further comprises mapping each of the at least two segments to a different predefined encoder model (col.42, ln.47-61, Lee discloses that each video object has an arbitrary shape, and that each video object is predefined according to its shape, thus, each video object or video portion is routed or assigned a predefined encoder model by a mask of alpha values or a binary mask; fig.35, note frame 1538 consists of multiple portions 1540, 1542, 1544a and 1544b; fig.33 and col.42, ln.62-65; note different video object coders 1504-1508 encode video portions associated with the generic model; col.43, ln.10-15; note the multiplexer 1510 is used to multiplex and encode video portions from plural different video object encoders 1504-1508).

Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Allen Wong whose telephone number is (571) 272-7341. The examiner can normally be reached on Mondays to Thursdays from 8am-6pm Flextime.

If attempts to reach the examiner by telephone are unsuccessful, the examiner’s supervisor, John W. Miller can be reached on (571) 272-7353. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Allen Wong/
Primary Examiner
Art Unit 2621

/Allen Wong/
Primary Examiner, Art Unit 2621
7/10/08